



OIEP

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/945,301

DATE: 01/27/2002  
 TIME: 15:35:35

Input Set : D:\38155-20036.txt  
 Output Set: N:\CRF3\01272002\I945301.raw

ENTERED

4 <110> APPLICANT: Millennium Pharmaceuticals, Inc.  
 5 Glucksmann, Maria  
 6 Tsai, Fong-Ying  
 8 <120> TITLE OF INVENTION: 27439, NOVEL HUMAN HYDROXYLASE AND USES  
 9 THEREFOR  
 11 <130> FILE REFERENCE: 38155-20036.00  
 13 <140> CURRENT APPLICATION NUMBER: US 09/945,301  
 14 <141> CURRENT FILING DATE: 2001-08-31  
 16 <150> PRIOR APPLICATION NUMBER: US 60/229,301  
 17 <151> PRIOR FILING DATE: 2000-09-01  
 19 <160> NUMBER OF SEQ ID NOS: 16  
 21 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 23 <210> SEQ ID NO: 1  
 24 <211> LENGTH: 1976  
 25 <212> TYPE: DNA  
 26 <213> ORGANISM: Homo sapiens  
 28 <220> FEATURE:  
 29 <221> NAME/KEY: CDS  
 30 <222> LOCATION: (1)...(1614)  
 32 <221> NAME/KEY: misc\_feature  
 33 <222> LOCATION: (1)...(1976)  
 34 <223> OTHER INFORMATION: n = A,T,C or G  
 36 <400> SEQUENCE: 1  
 37 atg gcc ctg ctg gcg cgg atc ctg aga gcc ggg ctg cgg ccg gcg ccc 48  
 38 Met Ala Leu Leu Ala Arg Ile Leu Arg Ala Gly Leu Arg Pro Ala Pro  
 39 1 5 10 15  
 41 gag cgg ggt ggg ctc ctg ggc ggc ggg gcc ccg cgg cgg cct caa ccc 96  
 42 Glu Arg Gly Gly Leu Leu Gly Gly Gly Ala Pro Arg Arg Pro Gln Pro  
 43 20 25 30  
 45 gcg ggc gca cgg ctc ccg gcg ggg gcg cgg gcc gag gac aaa ggc gcc 144  
 46 Ala Gly Ala Arg Leu Pro Ala Gly Ala Arg Ala Glu Asp Lys Gly Ala  
 47 35 40 45  
 49 ggg cgg ccg ggg tgc ccg ccg gga ggg ggc cga gcc gag ggt ccc cgg 192  
 50 Gly Arg Pro Gly Ser Pro Pro Gly Gly Gly Arg Ala Glu Gly Pro Arg  
 51 50 55 60  
 53 agc ctc gcc gcc atg ccg ggg ccg agg acc ctc gcc aac ctg gcg gag 240  
 54 Ser Leu Ala Ala Met Pro Gly Pro Arg Thr Leu Ala Asn Leu Ala Glu  
 55 65 70 75 80  
 57 ttc ttc tgc agg gac ggc ttc agc cgc atc cac gag atc cag cag aag 288  
 58 Phe Phe Cys Arg Asp Gly Phe Ser Arg Ile His Glu Ile Gln Gln Lys  
 59 85 90 95  
 61 cac aca cgg gaa tat gga aaa atc ttc aag tct cac ttt ggt cct cag 336  
 62 His Thr Arg Glu Tyr Gly Lys Ile Phe Lys Ser His Phe Gly Pro Gln

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/945,301

DATE: 01/27/2002

TIME: 15:35:35

Input Set : D:\38155-20036.txt

Output Set: N:\CRF3\01272002\I945301.raw

63	100	105	110	
65	ttt gta gta tct att gca gac cgc gat atg gtg gct cag gtg ctc cgg	384		
66	Phe Val Val Ser Ile Ala Asp Arg Asp Met Val Ala Gln Val Leu Arg			
67	115 120 125			
69	gcg gag ggc gct gcg ccc cag aga gcc aac atg gag tcc tgg cgg gag	432		
70	Ala Glu Gly Ala Ala Pro Gln Arg Ala Asn Met Glu Ser Trp Arg Glu			
71	130 135 140			
73	tac cga gac ttg cgg ggg aga gcc acc ggg ctc atc tcg gcg gag ggt	480		
74	Tyr Arg Asp Leu Arg Gly Arg Ala Thr Gly Leu Ile Ser Ala Glu Gly			
75	145 150 155 160			
77	gaa cag tgg ctc aag atg aga agc gta ttg aga caa aga att ctg aaa	528		
78	Glu Gln Trp Leu Lys Met Arg Ser Val Leu Arg Gln Arg Ile Leu Lys			
79	165 170 175			
81	ccg aaa gat gtg gcc att tat tct gga gaa gtc gac caa gtt att gct	576		
82	Pro Lys Asp Val Ala Ile Tyr Ser Gly Glu Val Asp Gln Val Ile Ala			
83	180 185 190			
85	gac tta att aaa aga atc tac ctc ctc agg agc cag gca gaa gat gga	624		
86	Asp Leu Ile Lys Arg Ile Tyr Leu Leu Arg Ser Gln Ala Glu Asp Gly			
87	195 200 205			
89	gaa acc gtg acc aat gtc aat gat ctt ttc ttc aaa tat tca atg gaa	672		
90	Glu Thr Val Thr Asn Val Asn Asp Leu Phe Phe Lys Tyr Ser Met Glu			
91	210 215 220			
93	gga gtg gcc acc atc ctt tat gag agt cgt ttg ggc tgc ctg gaa aac	720		
94	Gly Val Ala Thr Ile Leu Tyr Glu Ser Arg Leu Gly Cys Leu Glu Asn			
95	225 230 235 240			
97	agc atc cca cag ctg act gtg gaa tac atc gag gcc ctg gag ctc atg	768		
98	Ser Ile Pro Gln Leu Thr Val Glu Tyr Ile Glu Ala Leu Glu Leu Met			
99	245 250 255			
101	ttt agc atg ttc aag acc tcc atg tat gca ggc gcc atc ccc aga tgg	816		
102	Phe Ser Met Phe Lys Thr Ser Met Tyr Ala Gly Ala Ile Pro Arg Trp			
103	260 265 270			
105	ctt cgc ccc ttc atc cca aag ccc tgg cgg gaa ttc tgc agg tcc tgg	864		
106	Leu Arg Pro Phe Ile Pro Lys Pro Trp Arg Glu Phe Cys Arg Ser Trp			
107	275 280 285			
109	gat gga ctc ttc aaa ttc agc caa att cat gtt gac aac aag ttg tgg	912		
110	Asp Gly Leu Phe Lys Phe Ser Gln Ile His Val Asp Asn Lys Leu Trp			
111	290 295 300			
113	gac ata cag tac caa atg gac cga ggc cgg agg gtg agc ggg gga ctt	960		
114	Asp Ile Gln Tyr Gln Met Asp Arg Gly Arg Arg Val Ser Gly Gly Leu			
115	305 310 315 320			
117	ctc aca tac ctc ttc ctt agc cag gct ctg acg ctg cag gag atc tac	1008		
118	Leu Thr Tyr Leu Phe Leu Ser Gln Ala Leu Thr Leu Gln Glu Ile Tyr			
119	325 330 335			
121	gcc aac gtg act gag atg ctg ctg gcc ggc gtc gac acg acg tcc ttc	1056		
122	Ala Asn Val Thr Glu Met Leu Leu Ala Gly Val Asp Thr Thr Ser Phe			
123	340 345 350			
125	acc ttg tct tgg acg gtg tac ctc ctg gca agg cac cca gaa gtg cag	1104		
126	Thr Leu Ser Trp Thr Val Tyr Leu Leu Ala Arg His Pro Glu Val Gln			
127	355 360 365			

## RAW SEQUENCE LISTING

DATE: 01/27/2002

PATENT APPLICATION: US/09/945,301

TIME: 15:35:35

Input Set : D:\38155-20036.txt

Output Set: N:\CRF3\01272002\I945301.raw

```

129 cag acg gtg tac cgg gag att gtg aag aat tta ggg gaa agg cat gtt      1152
130 Gln Thr Val Tyr Arg Glu Ile Val Lys Asn Leu Gly Glu Arg His Val
131      370      375      380
133 cca act gca gct gat gtc ccc aag gtc ccg ctg gtc aga gct ctc ctt      1200
134 Pro Thr Ala Ala Asp Val Pro Lys Val Pro Leu Val Arg Ala Leu Leu
135 385      390      395      400
137 aag gaa acc ctg agg ctg ttt cca gtg ctg cca ggg aac ggc cgg gtc      1248
138 Lys Glu Thr Leu Arg Leu Phe Pro Val Leu Pro Gly Asn Gly Arg Val
139      405      410      415
141 acc cag gaa gac ctg gtt att ggc ggg tat ctg att ccg aaa ggc acc      1296
142 Thr Gln Glu Asp Leu Val Ile Gly Gly Tyr Leu Ile Pro Lys Gly Thr
143      420      425      430
145 cag ctg gcc ctt tgc cac tat gcc aca tcg cac cag gat gag aac ttc      1344
146 Gln Leu Ala Leu Cys His Tyr Ala Thr Ser His Gln Asp Glu Asn Phe
147      435      440      445
149 cct cgg gcc aag gag ttc cga cct gag cgc tgg ctg cgg aaa gga gac      1392
150 Pro Arg Ala Lys Glu Phe Arg Pro Glu Arg Trp Leu Arg Lys Gly Asp
151      450      455      460
153 tta gat aga gtt gac aat ttt gga tcc atc ccc ttt ggt cat ggg gtt      1440
154 Leu Asp Arg Val Asp Asn Phe Gly Ser Ile Pro Phe Gly His Gly Val
155 465      470      475      480
157 cgc agc tgc ata ggg cgg aga att gca gaa ctg gag att cac ctc gtc      1488
158 Arg Ser Cys Ile Gly Arg Arg Ile Ala Glu Leu Glu Ile His Leu Val
159      485      490      495
161 gtg atc cag ttg ctt caa cat ttt gag atc aaa aca tct tct cag acc      1536
162 Val Ile Gln Leu Leu Gln His Phe Glu Ile Lys Thr Ser Ser Gln Thr
163      500      505      510
165 gat gct gtt cat gca aaa acc cac ggg ctc ctg acg cca ggg ggg ccc      1584
166 Asn Ala Val His Ala Lys Thr His Gly Leu Leu Thr Pro Gly Gly Pro
167      515      520      525
169 atc cac gtg cga ttt gtt aac aga aag taa gcctagattt taaacctggg      1634
170 Ile His Val Arg Phe Val Asn Arg Lys *
171      530      535
173 ctgatgtagc agaccagctc gccgacacac agtgggtatt tgtgttcgct gatcaccgtg      1694
174 gagaaggaaa gcgatgtcgc taaaggctgt cttgttatag actggcctcc caggtcctgg      1754
175 gacacttgta aatctttatg caaagtaatg taaaaagggt gctattttac tgggtgcatac      1814
176 cagaagttgc cctttctttg ggggaaacag ctgtttaaaa accagtggca gtgaattttt      1874
177 atgcttcata cattngcta gactcaatat ttaatgactg gcagtatcct gtgcatttac      1934
178 ttgtacaggg aaatggtggn ttacttacaa attcagttct tc      1976
180 <210> SEQ ID NO: 2
181 <211> LENGTH: 537
182 <212> TYPE: PRT
183 <213> ORGANISM: Homo sapiens
185 <400> SEQUENCE: 2
186 Met Ala Leu Leu Ala Arg Ile Leu Arg Ala Gly Leu Arg Pro Ala Pro
187 1      5      10      15
188 Glu Arg Gly Gly Leu Leu Gly Gly Gly Ala Pro Arg Arg Pro Gln Pro
189      20      25      30
190 Ala Gly Ala Arg Leu Pro Ala Gly Ala Arg Ala Glu Asp Lys Gly Ala

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/945,301

DATE: 01/27/2002

TIME: 15:35:35

Input Set : D:\38155-20036.txt

Output Set: N:\CRF3\01272002\I945301.raw

```

191          35          40          45
192 Gly Arg Pro Gly Ser Pro Pro Gly Gly Gly Arg Ala Glu Gly Pro Arg
193          50          55          60
194 Ser Leu Ala Ala Met Pro Gly Pro Arg Thr Leu Ala Asn Leu Ala Glu
195 65          70          75          80
196 Phe Phe Cys Arg Asp Gly Phe Ser Arg Ile His Glu Ile Gln Gln Lys
197          85          90          95
198 His Thr Arg Glu Tyr Gly Lys Ile Phe Lys Ser His Phe Gly Pro Gln
199          100          105          110
200 Phe Val Val Ser Ile Ala Asp Arg Asp Met Val Ala Gln Val Leu Arg
201          115          120          125
202 Ala Glu Gly Ala Ala Pro Gln Arg Ala Asn Met Glu Ser Trp Arg Glu
203          130          135          140
204 Tyr Arg Asp Leu Arg Gly Arg Ala Thr Gly Leu Ile Ser Ala Glu Gly
205 145          150          155          160
206 Glu Gln Trp Leu Lys Met Arg Ser Val Leu Arg Gln Arg Ile Leu Lys
207          165          170          175
208 Pro Lys Asp Val Ala Ile Tyr Ser Gly Glu Val Asp Gln Val Ile Ala
209          180          185          190
210 Asp Leu Ile Lys Arg Ile Tyr Leu Leu Arg Ser Gln Ala Glu Asp Gly
211          195          200          205
212 Glu Thr Val Thr Asn Val Asn Asp Leu Phe Phe Lys Tyr Ser Met Glu
213          210          215          220
214 Gly Val Ala Thr Ile Leu Tyr Glu Ser Arg Leu Gly Cys Leu Glu Asn
215 225          230          235          240
216 Ser Ile Pro Gln Leu Thr Val Glu Tyr Ile Glu Ala Leu Glu Leu Met
217          245          250          255
218 Phe Ser Met Phe Lys Thr Ser Met Tyr Ala Gly Ala Ile Pro Arg Trp
219          260          265          270
220 Leu Arg Pro Phe Ile Pro Lys Pro Trp Arg Glu Phe Cys Arg Ser Trp
221          275          280          285
222 Asp Gly Leu Phe Lys Phe Ser Gln Ile His Val Asp Asn Lys Leu Trp
223          290          295          300
224 Asp Ile Gln Tyr Gln Met Asp Arg Gly Arg Arg Val Ser Gly Gly Leu
225 305          310          315          320
226 Leu Thr Tyr Leu Phe Leu Ser Gln Ala Leu Thr Leu Gln Glu Ile Tyr
227          325          330          335
228 Ala Asn Val Thr Glu Met Leu Leu Ala Gly Val Asp Thr Thr Ser Phe
229          340          345          350
230 Thr Leu Ser Trp Thr Val Tyr Leu Leu Ala Arg His Pro Glu Val Gln
231          355          360          365
232 Gln Thr Val Tyr Arg Glu Ile Val Lys Asn Leu Gly Glu Arg His Val
233          370          375          380
234 Pro Thr Ala Ala Asp Val Pro Lys Val Pro Leu Val Arg Ala Leu Leu
235 385          390          395          400
236 Lys Glu Thr Leu Arg Leu Phe Pro Val Leu Pro Gly Asn Gly Arg Val
237          405          410          415
238 Thr Gln Glu Asp Leu Val Ile Gly Gly Tyr Leu Ile Pro Lys Gly Thr
239          420          425          430

```

## RAW SEQUENCE LISTING

DATE: 01/27/2002

PATENT APPLICATION: US/09/945,301

TIME: 15:35:35

Input Set : D:\38155-20036.txt

Output Set: N:\CRF3\01272002\I945301.raw

```

240 Gln Leu Ala Leu Cys His Tyr Ala Thr Ser His Gln Asp Glu Asn Phe
241          435          440          445
242 Pro Arg Ala Lys Glu Phe Arg Pro Glu Arg Trp Leu Arg Lys Gly Asp
243          450          455          460
244 Leu Asp Arg Val Asp Asn Phe Gly Ser Ile Pro Phe Gly His Gly Val
245 465          470          475          480
246 Arg Ser Cys Ile Gly Arg Arg Ile Ala Glu Leu Glu Ile His Leu Val
247          485          490          495
248 Val Ile Gln Leu Leu Gln His Phe Glu Ile Lys Thr Ser Ser Gln Thr
249          500          505          510
250 Asn Ala Val His Ala Lys Thr His Gly Leu Leu Thr Pro Gly Gly Pro
251          515          520          525
252 Ile His Val Arg Phe Val Asn Arg Lys
253          530          535
255 <210> SEQ ID NO: 3
256 <211> LENGTH: 1614
257 <212> TYPE: DNA
258 <213> ORGANISM: Homo sapiens
260 <400> SEQUENCE: 3
261 atggccctgc tggcgcgat cctgagagcc gggctgcggc cggcgcccga gcgggggtggg      60
262 ctccctgggcg gcgggggcccc gcggcgccct caaccgcgg gcgcacggct cccggcgggg      120
263 gcgcgggcccg aggacaaaagg cgccgggccc ccgggggtgc cgccgggagg gggccgagcc      180
264 gaggggtcccc ggagcctcgc cgccatgccg gggccgagga cctcgccaa cctggcgag      240
265 ttctttctgca gggacggctt cagccgcctc cagcagatcc agcagaagca cacacgggaa      300
266 tatggaaaaa tcttcaagtc tcacttttgt cctcagtttg tagtatctat tgcagaccgc      360
267 gatattggtg ctcagggtgct ccgggcgagg ggcgtgcgc cccagagagc caacatggag      420
268 tcttgccggg agtaccgaga cttgcggggg agagccaccg ggctcatctc ggcgaggggg      480
269 gaacagtggc tcaagatgag aagcgtattg agacaaagaa ttctgaaacc gaaagatgtg      540
270 gccatttatt ctggagaagt cgaccaagtt attgctgact taattaaaag aatctacctc      600
271 ctcaggagcc aggcagaaga tggagaaacc gtgaccaatg tcaatgatct tttcttcaaa      660
272 tattcaatgg aaggagtggc caccatccct tatgagagtc gtttgggctg cctggaaaac      720
273 agcatcccac agctgactgt ggaatacatc gaggccctgg agctcatgtt tagcatgttc      780
274 aagacctcca tgtatgcagg cgccatcccc agatggcttc gccccttcac cccaaagccc      840
275 tggcggggaat tctgcaggtc ctgggatgga ctcttcaaat tcagccaaat tcatgttgac      900
276 aacaagttgt gggacataca gtaccaaag gaccgaggcc ggagggtgag cgggggactt      960
277 ctacataacc tcttccttag ccaggctctg acgctgcagg agatctacgc caacgtgact      1020
278 gagatgctgc tggccggcgt cgacacgacg tccttcacct tgtcttgagc ggtgtacctc      1080
279 ctggcaagcg acccaaggt gcagcagacg gtgtaccggg agattgtgaa gaatttaggg      1140
280 gaaaggcatg ttccaactgc agctgatgtc cccaagggtc cgctggtcag agctctcctt      1200
281 aaggaaaacc tgaggctgtt tccagtgtg ccagggaacg gccgggtcac ccaggaagac      1260
282 ctggttattg gcgggtatct gattccgaaa ggcaccacgc tggccctttg ccactatgcc      1320
283 acatcgcaac aggatgagaa ctccctcggg gccaggagt tccgacctga gcgctggctg      1380
284 cggaaaggag acttagatag agttgacaat ttggatcca tcccctttgg tcatgggggt      1440
285 cgcagctgca tagggcgagg aattgcagaa ctggagattc acctcgctgt gatccagttg      1500
286 cttcaacatt ttgagatcaa aacatcttct cagaccaatg ctgttcatgc aaaaaccac      1560
287 gggctcctga cgccaggggg gcccatccac gtgcgatttg ttaacagaaa gtaa      1614
289 <210> SEQ ID NO: 4
290 <211> LENGTH: 496
291 <212> TYPE: PRT

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/945,301

DATE: 01/27/2002

TIME: 15:35:36

Input Set : D:\38155-20036.txt

Output Set: N:\CRF3\01272002\I945301.raw

L:177 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1

L:178 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1